DOCUMENT RESUME

ED 038 281 **SE 008 055**

AUTHOR	Beard, Jean
TITLE	Group Achievement Tests Developed for Two Basic Processes of AAAS Science--A Process Approach.
INSTITUTION	San Jose State Coll., Calif. Dept. of Natural Science.
PUB DATE	Mar 70
NOTE	9p.; Paper presented at Annual Meeting of the National Association for Research in Science Teaching (43rd, Minneapolis, Minn., March 5-8, 1970)
EDRS PRICE	EDRS Price MF-$0.25 HC-$0.55
DESCRIPTORS	*Achievement Tests, *Elementary School Science, *Evaluation, *Group Tests, *Test Construction
IDENTIFIERS	Science (A Process Approach)

ABSTRACT

The major purpose of this study was to determine whether science process achievement tests could be developed which could be administered to groups of primary grade pupils. Tests developed were a series of 35mm color slides which illustrated laboratory situations involving basic science processes. The slides were synchronized with a tape recording to provide instructions for the children. Each pupil indicated his answer to each question by marking his answer sheet as directed from the tape recording. Six samples of the Basic Science Process Tests (BSPT) were developed. Each BSPT item was validated by a critique jury. Data were collected on 854 pupils in grades 1, 2, and 3, half of whom had studied the 1967 edition of "Science - A Process Approach" for one year. Evidence from this study indicates that the test format used can be made to function effectively as a group test for primary grade children. Pupils reacted favorably to the BSPT and seemed to be able to respond appropriately to the format and to individual items. (BR)

GROUP ACHIEVEMENT TESTS DEVELOPED FOR TWO BASIC PROCESSES OF AAAS SCIENCE--A PROCESS APPROACH

Jean Beard
Department of Natural Science
San Jose State College

Primary Study

The major purpose of the study was to determine whether science process achievement tests could be developed which could be administered to groups of primary grade pupils. A particular test format was proposed to assess pupil science process achievement. In this format pupils viewed and heard each test item.

The tests developed were a series of 35 mm colored slides which illustrate laboratory situations involving basic science processes. The slides are in sequence with a synchronized tape recording which orally provides instructions for the children and states the problem to be considered, and changes the slides. Each pupil indicates his answer to each question by marking his own answer sheet as directed from the tape recording. This procedure attempted to assess the science process skills of groups of primary pupils who had yet to perfect their reading and writing skills.

A paper presented March 7, 1970 in Minneapolis, Minnesota at the Annual meeting of the National Association for Research in Science Teaching

Six samples of the Basic Science Process Tests (BSPT) format were developed. Each BSPT was constructed to accompany the objectives taught in one of the basic sciences processes in one part of the 1967 edition of Science--A Process Approach (Sci-APA).

Test Development Procedures

Since the tests proposed had no comparable equivalents the test design procedures used to develop and establish them were unique to the present study. The procedures listed below began in February, 1968 and were culminated by the assembly of six sample BSPTs by the end of August, 1968. These tests, one each in the Measuring and Classifying processes to accompany each of Parts A, B, and C of Sci-APA were given during September, 1968 to Wisconsin pupils in grades one, two and three.

The chronology of procedures listed below was followed in developing and administering the BSPTs:

1. Selection of representative Basic Science Processes.
2. Design of the answer sheet format.
3. Design and construction of test items.
4. Validation of test items.
5. Development of general instructions and sample items.
6. Pilot testing with students of appropriate minimum ages.
7. Pilot testing with students of appropriate minimum ages and experience in Sci-APA.
8. Assembly of test items into sample BSPTs.
9. Production of synchronized tape recordings.
10. Administration of sample BSPTs.
11. Scoring of BSPTs and analysis of their performances.
12. BSPT item analysis from test data only.
13. Analysis of BSPT item group performances.

Validity and Reliability. Each BSPT item was validated by a critique jury. Each juror judged whether the item met two criteria.

First, the item had to meet the objective from Sci-APA it was constructed to assess. Second, a usable item was judged likely to elicite appropriate information from children of the minimum expected ages. BSPT reliabilities were estimated by correlating the test and retest results from fall 1968 administrations. A criterion correlation coefficient of 0.70 was set.

Pupil use of materials. Materials were used with pupils at three stages. First, spring 1968 pilot work was done with 70 Corvallis, Oregon pupils in kindergarten and first grade to establish the general testing procedures and to develop the BSPT directions and sample items. Second, Portland, Oregon pupils who had studied Sci-APA from the 1965 edition for one year were given groups of BSPT items. The 105 kindergarten, 50 first grade and 102 second grade pupils tested in the spring of 1968 provided information about BSPT item functioning, time requirements for each item, and test-retest time intervals. Third, assembled BSPTs were tried in the fall of 1968 with 854 pupils in Park Falls, Phillips and Rhinelander, Wisconsin, half of whom had studied the 1967 edition of Sci-APA only the previous year. The groups of first, second and third grade pupils, given two BSPTs twice, provided test-retest scores for BSPT reliability estimations.

Data Analysis. Fall 1968 data were used for detailed item and sample BSPT analysis. Information reported in the complete study of each BSPT included:

1. Reliability estimates.
2. Test and retest data--range, mean and standard deviation.
3. BSPT item performance--percent of correct responses, item difficulty, item discrimination, and items eliciting extra marks or omits.
4. Percent of correct responses on both the test and retest.

Each of the above analyses was done separately for the population of pupils who had Sci-APA experience and the population that lacked this science experience. In first grade a division of the non-Sci-APA population considered those with and those without kindergarten experience for each of the above analysis categories.

Secondary Study

A second purpose of the study was to determine whether the BSPTs that were developed distinguished between pupil populations that did and did not have experience in Sci-APA. Data from the different populations that took the same BSPTs was compared to determine what differences were evident in their performances. The t-test was applied to the difference of BSPT means, and Chi-squares were calculated for each item and for items grouped by Sci-APA exercises to test for homogeniety of item means across populations.

Summary and Conclusions

Evidence from the present study indicated that the proposed test format can be made to function effectively as a group test for primary grade pupils. Fall testing of first, second and third

grade pupils proceeded efficiently with the six sample BSPTs which were administered twice to each class. The data produced from the two BSPTs which accompany Part A of Sci-APA show these tests to be reliable at the agreed upon level and to contain items which function effectively. Table I is a summary of the data and statistics for the three BSPTs which are indicative of the format success. Classifying B, included in the table, is not sufficiently reliable, but did meet some of the performance standards expected of objective tests. The BSPTs omitted from the table, Measuring B, Measuring C and Classifying C, did not attain the reliability standard established for them and did not distinguish between the pupil populations that took them.

Pupils reacted favorably to the BSPTs and seemed to be able to respond appropriately to the format and to individual items. The test lengths were apparently reasonable for first administration.

Several features of the BSPTs recommend them over other testing techniques. Groups of pupils saw and heard information in these tests and were able to indicate, individually and quietly, their decisions about the questions. All pupils had an equal opportunity to consider each item. The assembled BSPT once started continued to operate automatically from the signals on the tape recording, leaving the administrator free to observe and assist students.

Table I. Summary of Fall 1968 Data Supporting BSPT Format

Characteristic	Measuring A	Classifying A	Classifying B
Grade	1	1	2
Number of Items	9	12	16
Test-Retest Reliability Estimate			
AAAS	.70	.71	.50
X AAAS, Kg	.60	.50	.35
X AAAS, X Kg	.70	.62	---
Test Mean			
AAAS	6.74	7.62	10.01
X AAAS, Kg	6.59	7.06	9.16
X AAAS, X Kg	5.00	5.80	---
Expected Mean	5.65	7.69	9.96
t-test on Mean Difference			
AAAS X AAAS	.01	.01	.05
X AAAS, Kg X Kg	.01	.01	---
Number of Items That Discriminated Between Populations as Indicated by Chi-square Test for Homogeneity of Item Means	7	7	8

Two factors seemed to contribute to the limited success of the BSPTs built to accompany the second and third parts of Sci-APA. First, the pupils at all grade levels had studied Sci-APA for a maximum of one year, starting with Part A, B, or C. Thus, the second and third grade pupils had one-half or one-third of the recommended time in Sci-APA.

The second factor was related to the increasingly complex science content introduced at succeeding levels of a process hierarchy. The process skills taught at each succeeding level are less common to learning in general, and more specifically science oriented. Thus process skills taught in Parts B and C of Sci-APA would less frequently be those taught in other subject matter contexts.

The two BSPTs which were reliable distinguised between pupils who did and who did not have Sci-APA experience. The success of the Part A BSPTs is of particular interest because these were the shortest tests, administered to the youngest pupils in the study. Future trials of the BSPTs constructed to accompany the second and third parts of Sci-APA are expected to yield more reliable scores and discriminate between populations when used to assess pupils who have studied Sci-APA for two or three years.

This study seems to indicate that appropriate testing can be extended to groups of students that have limited reading and writing skills. Multimedia tests can gather information about

performance abilities of many age groups and in a variety of subject areas. Other media combinations will also be effective. The BSPTs produced in this study can easily be converted to 16 mm color sound film, and thus function with equipment available in all schools to yield tests administrable by classroom teachers.

If science process teaching is to be directed toward skill and concept mastery, as the Sci-APA process hierarchies suggest, criterion testing is called for. Power tests, such as those produced in this study, can easily be criterion tests. Such multimedia tests will easily fit into self-learning centers, thus being useful for individual as well as group assessment.

In order to produce more appropriate tests for assessing science achievement, such as those developed in this study, a greater investment of time, effort and money will be required than for today's standardized tests. If science teaching is moving toward more student involvement with representative natural phenomena shouldn't science testing be less abstract?

REFERENCES

American Association for the Advancement of Science. Commission on Science Education. 1967-1968. Science--a process approach. Parts A-E. New York, Xerox. Various paging (Boxed booklets)

_____ 1967. Science--a process approach/hierarchy chart. New York, Xerox. 1 sheet.

American Psychological Association. 1966. Standards for educational and psychological tests and manuals. Washington, D.C. 40 p.

Beard, Jean. 1970. Group achievement tests developed for two basic processes of AAAS Science--a process approach. Ed.D. thesis. Corvallis, Oregon State University. 149 numb. leaves.

Ebel, Robert L. 1965. Measuring educational achievement. Englewood Cliffs, N.J., Prentice-Hall. 481 p.

Gagne, Robert M. 1967. Science--a process approach, purposes, accomplishments, expectations. Washington, D.C., American Association for the Advancement of Science. 17 p.

Jacobs, James N. and Joseph L. Felix. 1968. Testing and educational and psychological development of preadolescent children--ages 6-12. Review of Educational Research 38:19-28.

Lindquist, E. F. (ed.). 1951. Educational measurement. Washington, D.C., American Council on Education. 819 p.

Smith, Herbert A. 1969. Curriculum development and instructional materials. Review of Educational Research 39:397-413.

Welch, Wayne W. 1969. Curriculum Evaluation. Review of Educational Research 39:429-443.